# bkgoptrate

February 23, 2011

**Abstract**

This task calculates the optimum ratio (or count), based on the given (background) time-series FITS and prints it to STDOUT.

## 1 Instruments/Modes

**bkgoptrate** is not XMM-specific: it can be applied to any FITS file.

## 2 Use

| | |
|---|---|
| pipeline processing | yes |
| interactive analysis | yes |

## 3 Description

In general any data include the time-variable background. If the background during a certain and limited period in the observation is too high compared to the averaged ratio of the intensity of the source to background, then those period with high background had better be excluded from the data. As the background level sometimes flares up during X-ray observations, we call it background (BKG) flare-filtering.

Now it is not a trivial question what is the optimum level of the BKG-flare-filtering.

### 3.1 Theory

A possible and good algorithm to achieve the objective, which this task follows, is to search for the point, at which the maximum signal-to-noise (S/N) ratio is achieved for the given background time-series after the bins above a threshold are excluded.

The S/N ratio is described as

$$\mathrm{S/N} = \frac{\mathrm{SourceCount}}{\sqrt{\mathrm{BackgroundCount}}}, \quad (1)$$

providing the standard Poisson statistics are applicable.

Let us define $Z^*$ as the reference threshold count per time bin, and $\mathrm{Exp}^*$ as the exposure time for the bin. When we use $i$ subscripts to represent each time bin, the above formula is rewritten as

$$\mathrm{S/N} = \frac{\Sigma_{i(\mathrm{Cnt}>Z_i)}\mathrm{Exp}_i \times \mathrm{SourceRate}}{\sqrt{\Sigma_{i(\mathrm{Cnt}>Z_i)}\mathrm{Cnt}_i}} \quad (2)$$

$$Z_i \equiv Z^* \times \frac{\mathrm{Exp}_i}{\mathrm{Exp}^*}, \quad (3)$$

$$(4)$$

where SourceRate, $\mathrm{Exp}_i$, $\mathrm{Cnt}_i$ and $Z_i$ are the count rate of the source, exposure, count of the given time-series and threshold count in each bin $i$, respectively.

What we here search for is the optimum $Z^*$, which is the background level, with which S/N becomes the maximum. In principle it requires a minimization (or maximization) method for search. We assume SourceRate is constant, then we can ignore SourceRate in comparing S/N ratios for different thresholds ($Z^*$):

$$\mathrm{S/N} \propto \frac{\Sigma_{i(\mathrm{Cnt}>Z_i)}\mathrm{Exp}_i}{\sqrt{\Sigma_{i(\mathrm{Cnt}>Z_i)}\mathrm{Cnt}_i}} \quad (5)$$

$$\mathrm{S/N} \propto \frac{\Sigma_{i(\mathrm{Cnt}>Z_i)}\mathrm{Exp}_i}{\sqrt{\Sigma_{i(\mathrm{Cnt}>Z_i)}\mathrm{Rate}_i \times \mathrm{Exp}_i}}, \quad (6)$$

where the second form represents the one with the rate instead of count of the input time-series, and $Z_i$ is defined in Eq. 3.

If all the time bins have the same exposure, then this formula is further simplified:

$$\mathrm{S/N} \propto \frac{\Sigma_{i(\mathrm{Cnt}>Z^*)}i}{\sqrt{\Sigma_{i(\mathrm{Cnt}>Z^*)}\mathrm{Cnt}_i}} \quad (7)$$

$$\mathrm{S/N} \propto \frac{\Sigma_{i(\mathrm{Rate}>Z^*)}i}{\sqrt{\Sigma_{i(\mathrm{Rate}>Z^*)}\mathrm{Rate}_i}}. \quad (8)$$

## 3.2 Practical behaviour and command-line parameters

This task reads the background time-series FITS-Table (`tssettabname`) for the specified X and Y columns (`xcol` and `ycol`) for TIME and RATE (or COUNT), respectively, searches for the optimum background count (or rate) per bin, then if found, print the value to standard output (`STDOUT`). In practice, all the bins for Time are assumed to be the same in size, that is, duration.

This task uses a very simple minimization method; it finds the bin(s) that has the highest rate then calculate the S/N with those bins excluded, and compares with the previous S/N, namely the S/N with those bins included. This procedure continues until the newly calculated S/N is lower than the the previous value, then it prints out the previous highest rate (NOT the count), or strictly speaking the value that is slightly higher than the previous highest rate (so that the users do not have to worry about whether to use "larger/smaller than" or "equal to or larger/smaller than" in the later processing). The assumption here is the S/N as a function of $Z^*$ is monotonic.

The unit of the time-series can be significant. Users can specify it with `tsstyle`=(auto—rate—count). If it is "auto" (Default), this task determines `tsstyle` (1) according to `ycol` if specified, (2) or if not, this task searches for the name of the column of RATE, COUNT and COUNTS in this order and determines `ycol` first, then `tsstyle` accordingly.

Although the size of each time bin is assumed to be the same, it can be defined in the different column in the table (in the same FITS extension as the X(time) and Y(count) axes), typically `FRACEXP`, which can be explicitly specified with `fracexpcol`.

Users can also control how this task treats the `FRACEXP` column with `fracexpstyle`=(auto—calc—threshold—none). If `fracexpstyle=none` this task simply ignores `FRACEXP` column and assumes all the bins have the same duration[1]. If `fracexpstyle=threshold`, then this task ignores the bins, of which `FRACEXP` is smaller than the user-specified parameter, `fracexplower`. If `fracexpstyle=calc`, then this task takes `FRACEXP` into account in the search. If `fracexpstyle=auto` (Default), then this task searches for the column `fracexpcol` (`FRACEXP` in default), and sets `fracexpstyle` to be (1) "calc" if found, else (2) "none".

Users can also specify the start and end times to consider in the search for the optimum background level with the parameters of `starttime` and `endtime`.

Users can also specify the bins to be ignored if the count (or rate) of the bin is lower than the value, `lowercutoffcount`. Note that the bins with negative count (or rate) or with undefined value are always ignored.

Users can also specify the ratio of the minimum time to be left where the further search stops, with `mintimeratio` (0.05 in default). Namely if the search for the optimum background cuts continues to the extent smaller than the fraction of `mintimeratio` (5% in default) in the exposure to the original exposure after all the other filtering is performed is left to be valid (in other words if over 95% in default of the exposure is regarded as high-background area), then this task stops further search for the optimum background level and instead raises an error (`noThresholdFound`).

Finally, if the large number of bins have the smallest and identical count, then chances are the maximum S/N is achieved when only those bins are used. For example, it is the case for the perfectly uniform time-series (although it must be an extremely rare case, if ever, with the real data). In that case, this task either (1) raises an error (`noThresholdFound`) preceded by a warning (`recommendedLowerMinTimeRatio`), or (2) returns the value slightly higher than the particular value preceded by a warning (`regardLowestBinsAsOptimum`) if all the time-bins are examined (due to a very small `mintimeratio` given) and if those time-bins occupy more than a third of the number of entire time-bins.

# 4 Parameters

This section documents the parameters recognized by this task (if any).

| **Parameter** | Mand | Type | Default | Constraints |
|---|---|---|---|---|
| **tssettabname** | yes | dataset | | |

Name of the input FITS data set and table, the latter of which can be added, following the ':', such as, 'input1.ds:EVENTS'.

| **xcol** | no | string | TIME | |
|---|---|---|---|---|

Column name of X-axis.

| **ycol** | no | string | default | |
|---|---|---|---|---|

Column name of Y-axis.

| **fracexpcol** | no | string | FRACEXP | |
|---|---|---|---|---|

Column name of FRACEXP.

[1] If a bin has an undefined value of count or rate, then the bin is ignored.

| | | | | |
|---|---|---|---|---|
| **tsstyle** | no | string | auto | auto—rate—count |

Style of the input time-series.

| | | | | |
|---|---|---|---|---|
| **withstarttime** | no | boolean | no | |

Whether start time is specified (Default: No). If yes, the task reads `starttime`.

| | | | | |
|---|---|---|---|---|
| **starttime** | yes | real | | |

Start time to evaluate.

| | | | | |
|---|---|---|---|---|
| **withendtime** | no | boolean | no | |

Whether end time is specified (Default: No). If yes, the task reads `endtime`.

| | | | | |
|---|---|---|---|---|
| **endtime** | yes | real | | |

End time to evaluate.

| | | | | |
|---|---|---|---|---|
| **fracexpstyle** | no | string | auto | auto—calc—threshold—none |

Style of handling method of FRACEXP. If `threshold`, the task reads `fracexplower`.

| | | | | |
|---|---|---|---|---|
| **fracexplower** | yes | real | | $0 \geq f \geq 1$ |

Lower threshold for FRACEXP to be used in the calculation.

| | | | | |
|---|---|---|---|---|
| **withlowercutoffcount** | no | boolean | no | |

Whether the lower cutoff count (or rate) is specified. If yes, the task reads `lowercutoffcount`.

| | | | | |
|---|---|---|---|---|
| **lowercutoffcount** | yes | real | 0.0 | $\geq 0$ |

The lower threshold count or ratio in search for the optimum value, that is, any bin having the count (or ratio) smaller than this value is not taken into account in the search. This value has to be either zero or positive.

| | | | | |
|---|---|---|---|---|
| **withmintimeratio** | no | boolean | no | |

Whether the minimum time ratio is specified. If yes, the task reads `mintimeratio`.

| | | | | |
|---|---|---|---|---|
| **mintimeratio** | yes | real | 0.05 | $\geq 0$ |

The ratio of the acceptable minimum number of (time) bins (or row numbers in the FITS table) to the initial number of bins after other filtering is applied, during the search for the optimum value. If the search does not find the optimum threshold with larger number of bins, then the task raises an error and exits.

# 5 Errors

This section documents warnings and errors generated by this task (if any). Note that warnings and errors can also be generated in the SAS infrastructure libraries, in which case they would not be documented here. Refer to the index of all errors and warnings available in the HTML version of the SAS documentation.

**noThresholdFound** *(error)*
Failed to find the optimum background (in the given constraints).

**regardLowestBinsAsOptimum** *(error)*

Regarded the lowest bins as the optimum background level (see text for detail).

**recommendedLowerMinTimeRatio** *(warning)*

*corrective action:* The lowest bins looks like the optimum background level in the given `mintimeratio` (see text for detail).

**negativeLowerCutoffY** *(warning)*

*corrective action:* Specified `lowercutoffcount` is negative.

# 6 Input Files

The input FITS needs not be XMM time-series but any FITS tables.

1. (Mandatory) `tssettabname`: the input FITS-table.

# 7 Output Files

Nil (all the outputs are printed to STDOUT).

# 8 Algorithm

# 9 Comments

•

# References